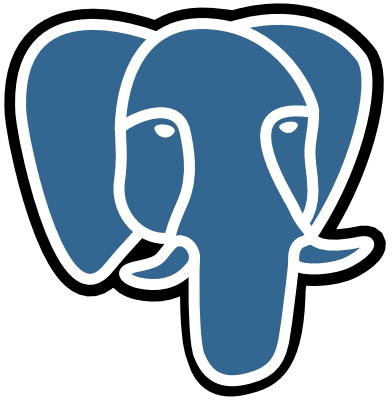
PostgreSQL

The world's most advanced open source database.

Czech and Slovak PostgreSQL Users Group



PostgreSQL for Oracle DBA

Oracle database logo should be there... but as of this page it can't be:
...Remember that you are generally not able to use Oracle logos unless you are specifically licensed or authorized to use them. ...

Aleš Zelený (zeleny.ales@gmail.com), CSPUG,

Agenda

- Mind Migration
- Some terminology
- “Architecture”
- Security
- Backup and Recovery
- High Availability / Disaster recovery
- Other unordered stuff to consider

Purpose of these slides

- Understand differences
- Commons in behavior
- Displeasure prevention

- Migration guide
- RDBMS ranking
- Feature matrix

Mind Migration

PostgreSQL
The world's most advanced open source database.

- PostgreSQL is not an Oracle database
- Oracle and PostgreSQL are both superb databases, their relation is not like Red Hat and CentOS
- Do not expect equivalents for all of the Oracle RDBMS features in PostgreSQL
- Don't hesitate to be impressed by PostgreSQL broad range of data types
- PostgreSQL has extensions

Terminology

- Architecture/concepts
  - Cluster
  - Instance
  - Database
  - Tablespace

Yes, a cluster can simply mean a grouping of related things, but context is everything.
Shaun M. Thomas

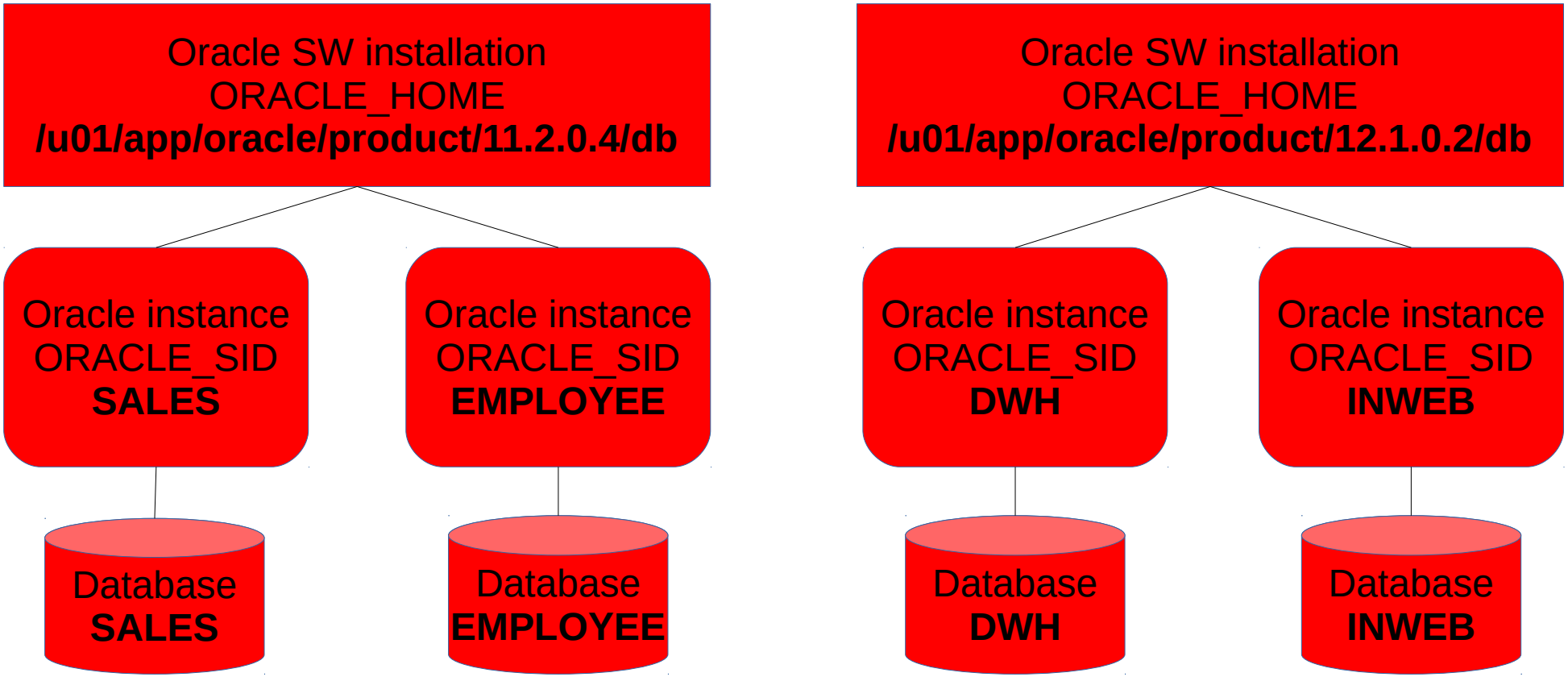
- Logical
  - Role
  - User
  - Schema
  - Tuple Wiki: A tuple is a finite ordered list of elements

- quick browse through manuals will help and don't take ages
  - Architectural Fundamentals
  - Documentation index

Oracle and DB relation

PostgreSQL
The world's most advanced open source database.

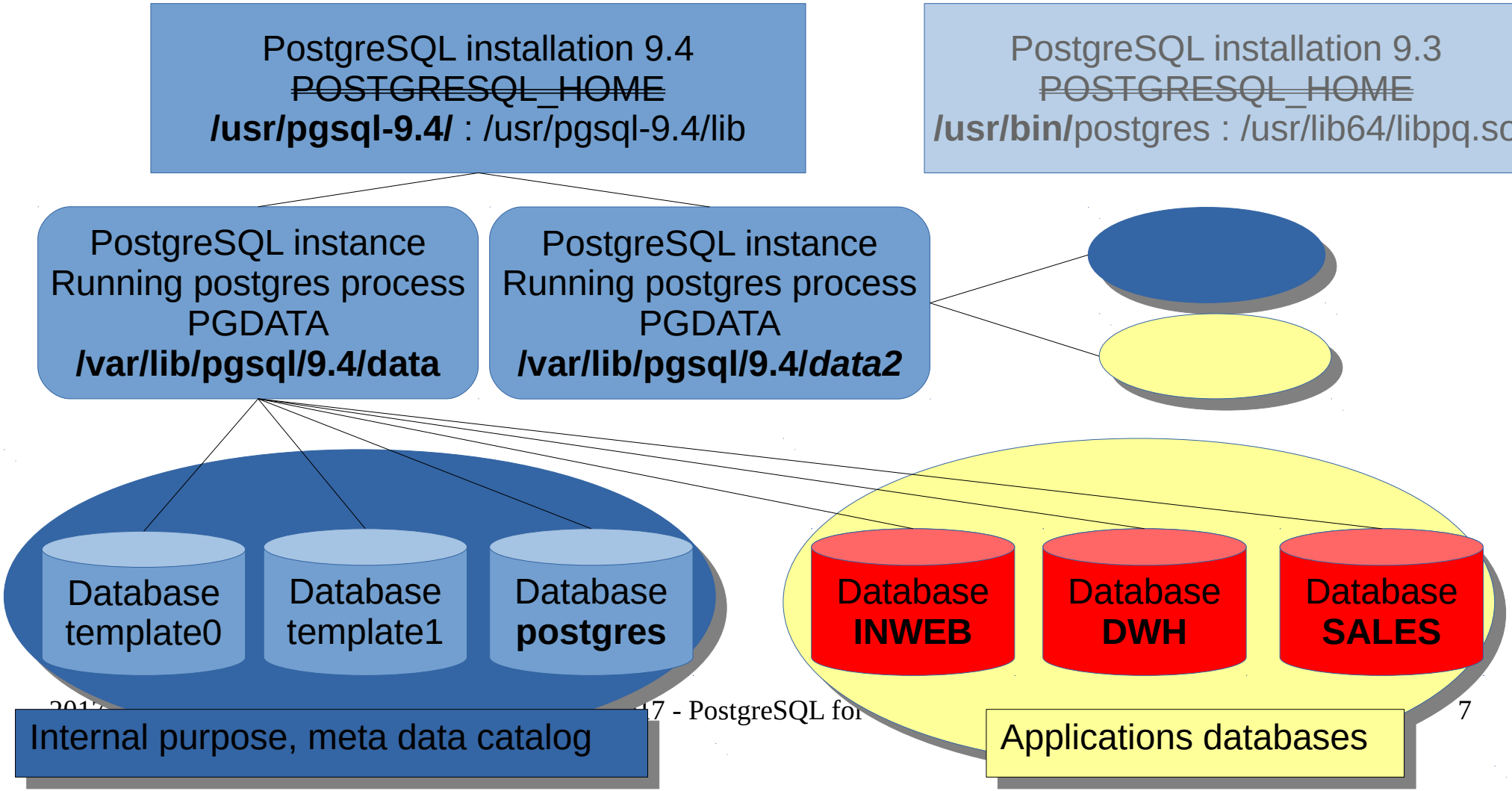
- From installed software to database (simplified)
  - Instance is software loaded into memory working with **ONE** database (12c PDB changed that rule)



PostgreSQL and DB relation

PostgreSQL
The world's most advanced open source database.

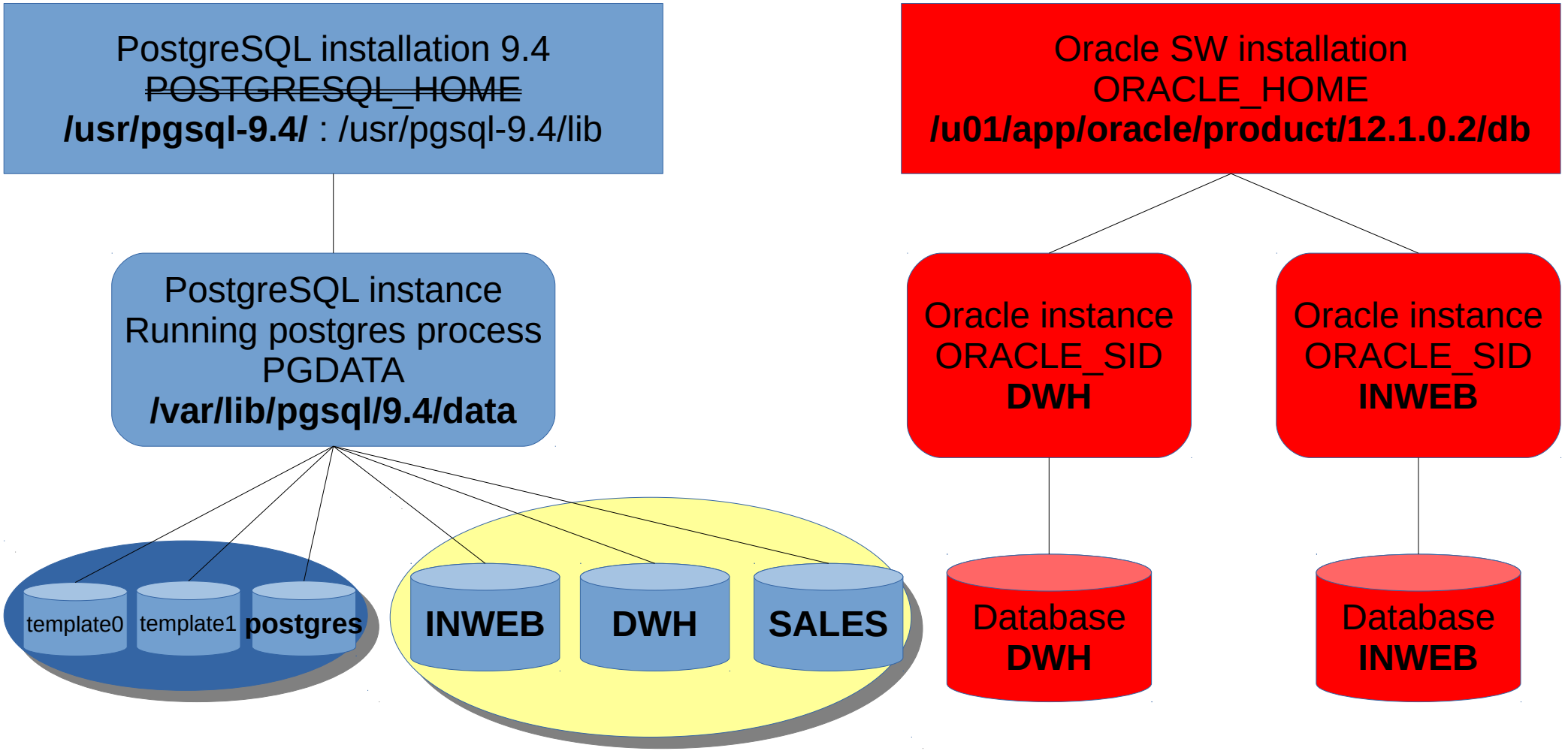
- From installed software to database (simplified)
  - SW installed from RPM/APT/compiled from source



Pg & Oracle – compared

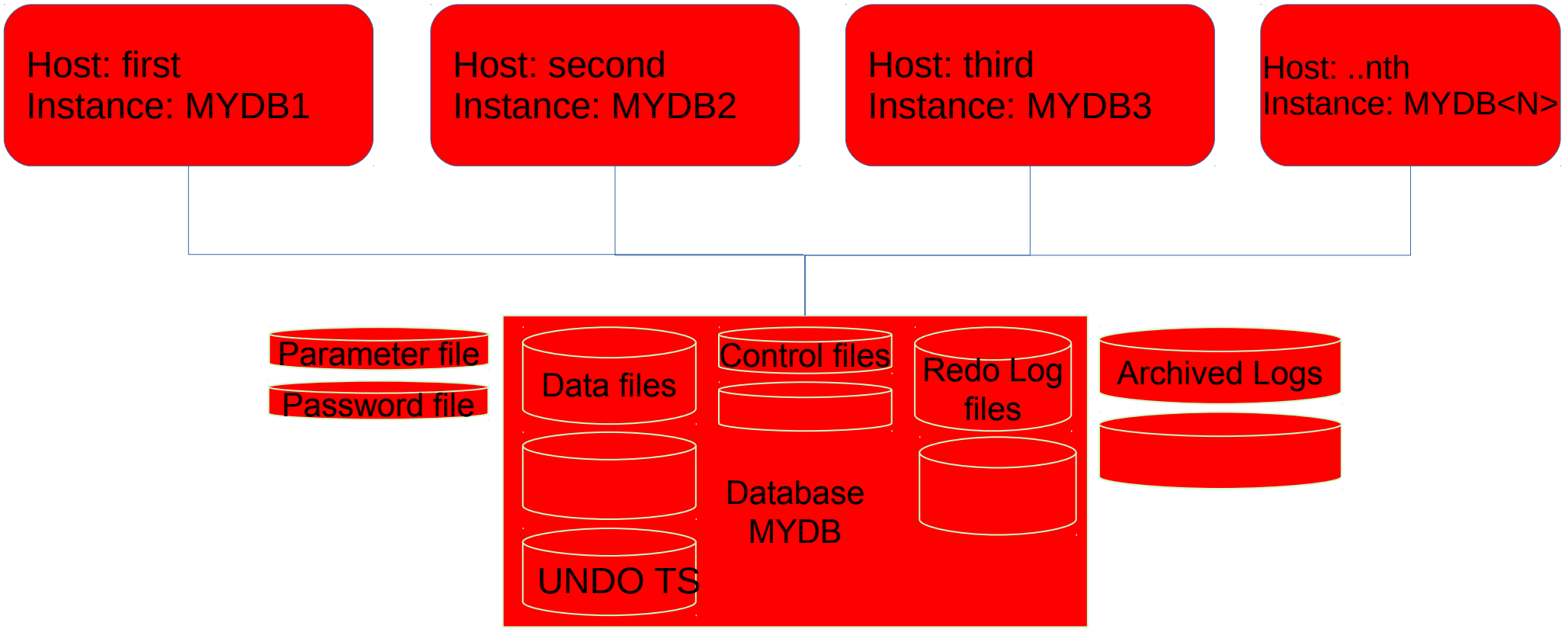
PostgreSQL
The world's most advanced open source database.

- From installed software to database (simplified)



Oracle – DB physical structure

- Oracle RAC (Real Application Clusters)
  - Even with RAC setup, an Oracle instance serves **ONE** database
    - Multitenant pluggable databases in version 12 breaks that rule...
- Oracle database physical components (files)



Postgres – cluster physical structure

PostgreSQL
The world's most advanced open source database.

Not means clustering like Oracle RAC

- Postgres uses directory (might be referred by environment variable PGDATA) traditionally called **database cluster** to store all necessary data managed by Postgres instance
  - The directory is mandatory parameter for starting Postgres instance (like [s]pfile for Oracle)
  - Contains configuration files, and in default setup also files for all databases residing within a particular Postgres cluster, see documentation Database File Layout

```
-bash-4.3$ ps -fu postgres
UID        PID  PPID  C STIME TTY          TIME CMD
postgres 30007     1  0 19:03 ?        00:00:00 /usr/bin/postgres -D /var/lib/pgsql/data -p 5433
```

database cluster

```
base/
global/
pg_clog/
pg_hba.conf
pg_ident.conf
pg_log/
pg_multixact/
pg_notify/
pg_serial/
pg_snapshots/
```

```
pg_stat/
pg_stat_tmp/
pg_subtrans/
pg_tblspc/
pg_twophase/
PG_VERSION
pg_xlog/
postgresql.conf
postmaster.opts
postmaster.pid
```

Configuration files

Databases directory

```
postgres=# select oid, datname from
pg_database;
  oid  |  datname
-------+-----------
     1 | template1
 12968 | template0
 12973 | postgres
(3 rows)
```

```
-bash-4.3$ du -sh base/*
6.4M    base/1
6.4M    base/12968
6.5M    base/12973
```

```
tree base | head -5
base
├── 1
│   ├── 12706
│   ├── 12706_fsm
│   ├── 12706_vm
```

Postgres – mapping to Oracle files

PostgreSQL
The world's most advanced open source database.

- Postgres similar to Oracle might create archive of online logs for PITR and other purposes
  - Instead of “Oracle Archiver” server processes, Postgres used to call any external command responsible for copying inactive online log to some other destination

```
-bash-4.3$ ps -fu postgres
UID        PID  PPID  C STIME TTY          TIME CMD
postgres 30007     1  0 19:03 ?        00:00:00 /usr/bin/postgres -D /var/lib/pgsql/data -p 5433
```

database cluster

```
base/
global/
pg_clog/
pg_hba.conf
pg_ident.conf
pg_log/
pg_multixact/
pg_notify/
pg_serial/
pg_snapshots/
```

sqlnet.ora
tcp.invited_nodes

```
pg_stat/
pg_stat_tmp/
pg_subtrans/
pg_tblspc/
pg_twophase/
PG_VERSION
pg_xlog/
postgresql.conf
postmaster.opts
postmaster.pid
```

Online Redo
Parameter file

Configuration files

Databases directory

```
postgres=# select oid, datname from
pg_database;
  oid  |  datname
-------+-----------
     1 | template1
 12968 | template0
 12973 | postgres
(3 rows)
```

```
tree base | head -5
base
├── 1
│   ├── 12706
│   ├── 12706_fsm
│   └── 12706_vm
```

Data files

Control files

Password file

archive_command ...
Archived Logs

documentation

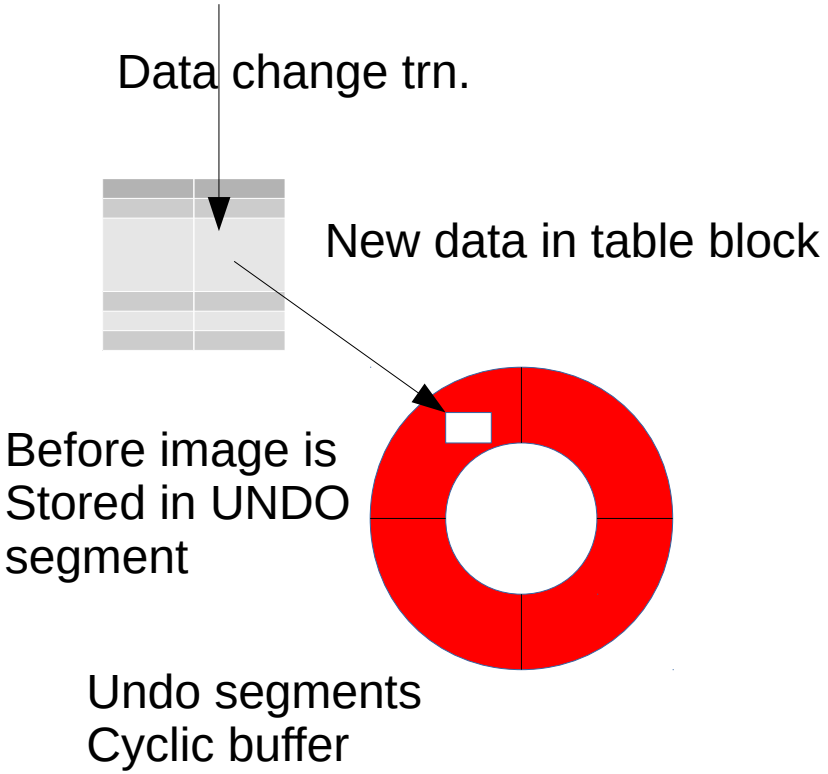
Tablespaces and filesystems

PostgreSQL
The world's most advanced open source database.

- Tablespaces might reside on different filesystem
  - Outage prevention (in case of full file system)
  - Data storage tiers, life cycle management
    - Online active data on SSD
    - Archive data on rotating disks
    - Tablespace for temporary files – fast might be unprotected storage, no data loss risk
- Wiki page - File System Layouts
- Tablespace is PG cluster level object, multiple databases might use it for some of their files

Where is UNDO tablespace?

- Answer: inside the data files
- Is this feature free of charge?
- No, space maintenance (vacuum) is needed to avoid table bloat.



- Postgres use MVCC model (Multiversion Concurrency Control) for data to transaction consistency
  - Transaction isolation for each session
  - Snapshot of data visible to each session based on transaction number
  - Minimize locking contention
  - Readers never blocks writes
  - Serializable snapshot Isolation is available

Online REDO ~ WAL files

PostgreSQL
The world's most advanced open source database.

- Online REDOlogs are cyclic buffer in Oracle
  - “cleaned up” by archiver process
  - Static amount of redolog Groups each with one or more members within a redolog group
  - Log switch tunning
- WAL – Write Ahead Log files (XLOGs)
  - “cyclic buffer space” with only soft limit in size
  - File reuse – rename already archived file
  - archive_command is used called each time WAL is switched to new file (there is no “archiver” process in postgres)

WAL files

- Place them on separate filesystem
  - Up to 9.4 space requirement for XLOG filesystem
    - **(2 + checkpoint_completion_target) * checkpoint_segments + 1** or **checkpoint_segments + wal_keep_segments + 1** files. Each segment file is normally 16 MB.
  - Starting with 9.5
    - **wal_min_size** (default 80MB ~ 5 xlogs) and **wal_max_size** (default 1GB ~ 64 xlogs)
  - Amount of WAL segments between automatic WAL checkpoint (higher values ~ potentially more data for datafiles recovery after server crash)
  - Both configurations are SOFT limit only

Archive_mode and WAL_level

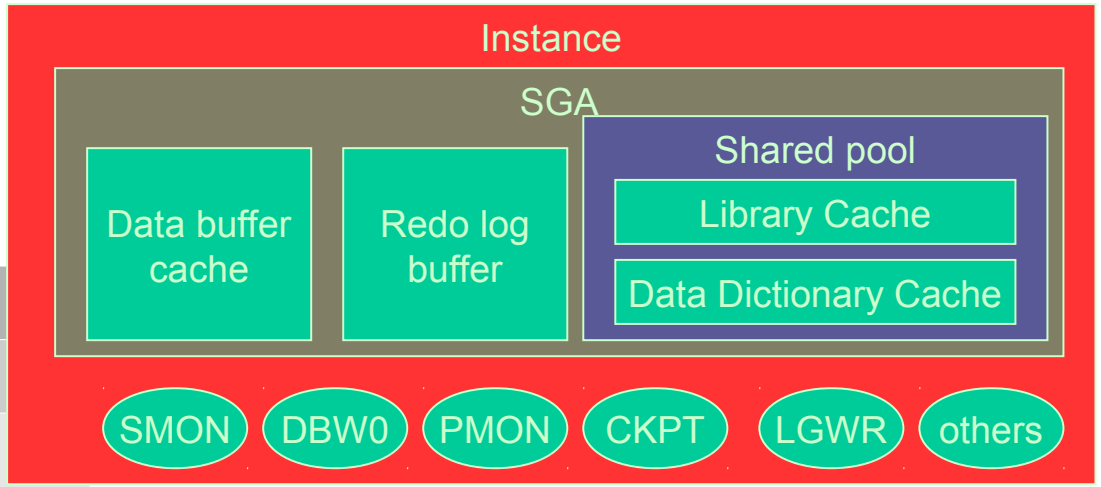
PostgreSQL
The world's most advanced open source database.

- Archive_mode
  - off, on, always (archive again on streaming replica)
- wal_level
  - minimal
    - Used for crash recovery only ~ Oracle noarchivelog
  - replica (from PG 9.6) ~ archive & hot_standby
    - archive (up to 9.5) – needed for Point In Time recovery
      - Used for replication (optionally streaming) ~ Oracle DataGuard archiver log shipping (Log Writer shipping)
    - hot_standby (up to 9.5)
      - Used for streaming replication with read only access to replica ~ Oracle Active DataGuard
  - logical
    - Supports logical decoding (~ Oracle logical standby)

Memory

PostgreSQL
The world's most advanced open source database.

Oracle structure:



Parameters relation….

Oracle	Postgres
db_cache_size	shared_buffers
sort_area_size (pga_aggregate_target)	work_mem, temp_buffers
log_buffer_size	wal_buffers
	maintetance_work_mem
	effective_cache_size

Postgres structure:

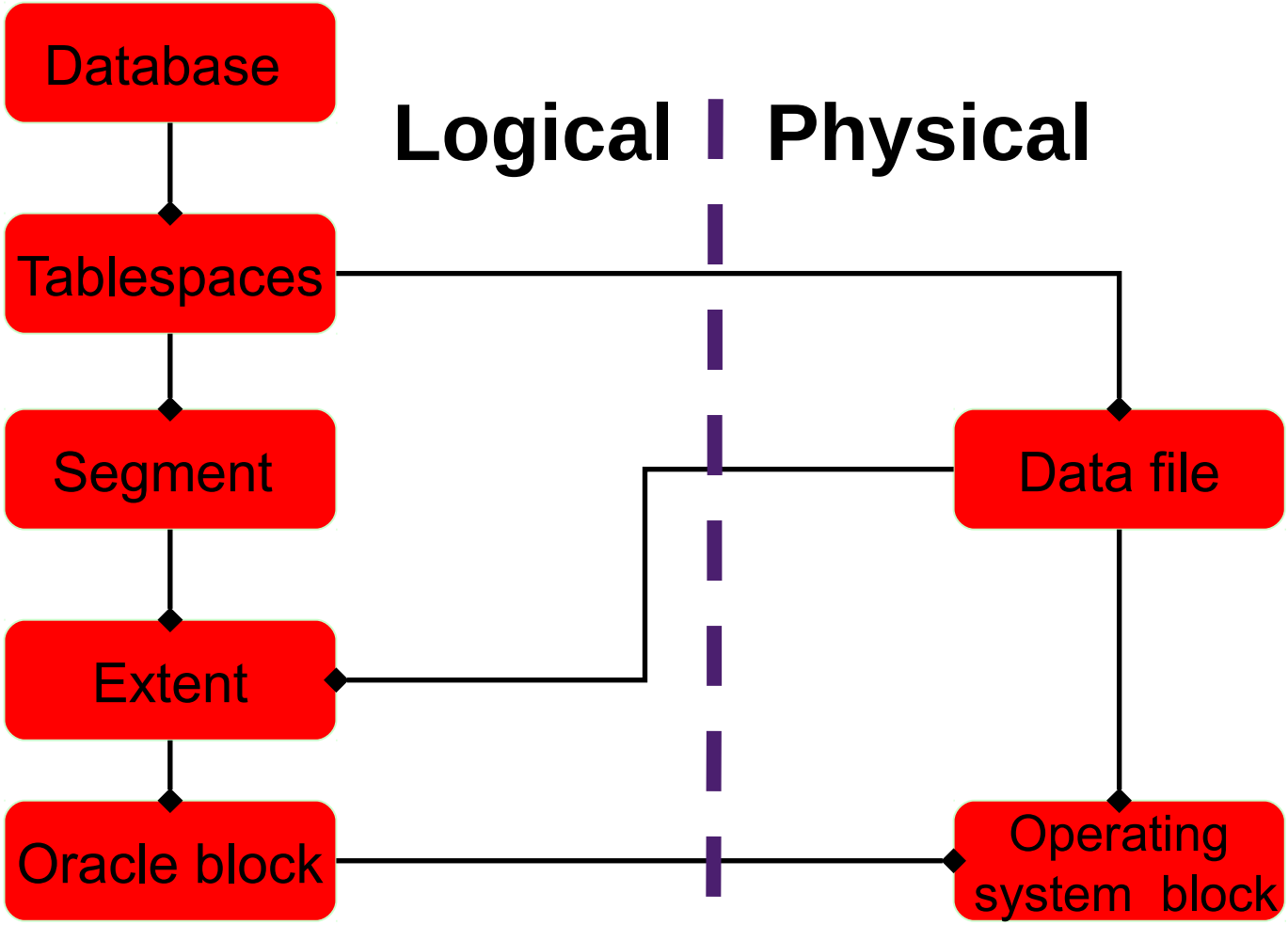
Nice description can be found at link:
PostgreSQL 9.0 Architecture

Per process MEMORY
- Work mem
- Maintenanace work mem
- Temp buffer
- Catalog cache
- Optimizer/executor

SHARED MEMORY
- shared buffers
- wal buffers
- CLOG buffers
- Locks space
- Other buffers

Architecture – database structure

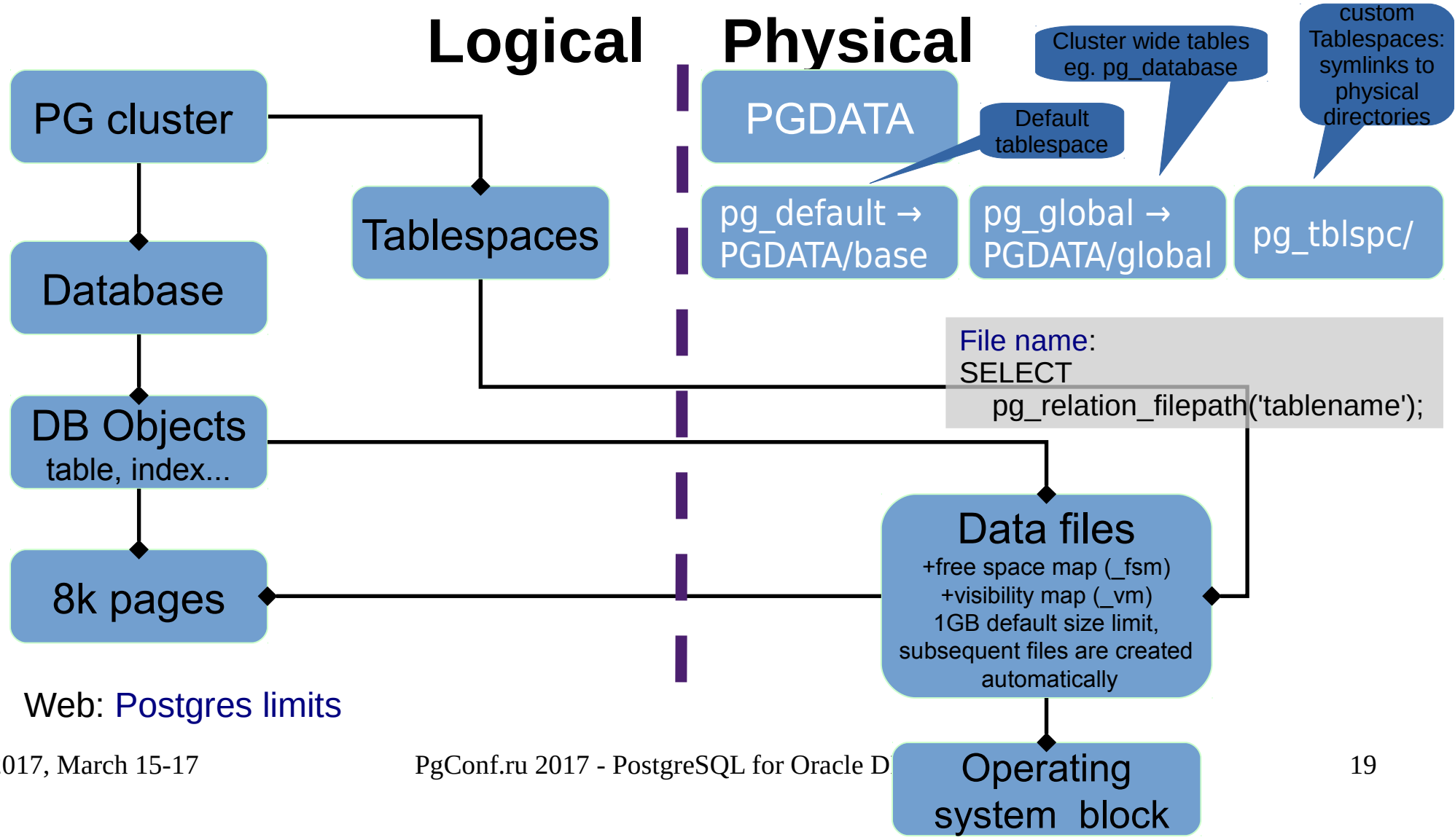
- Oracle database structure (simplified)



Architecture – database structure

PostgreSQL
The world's most advanced open source database.

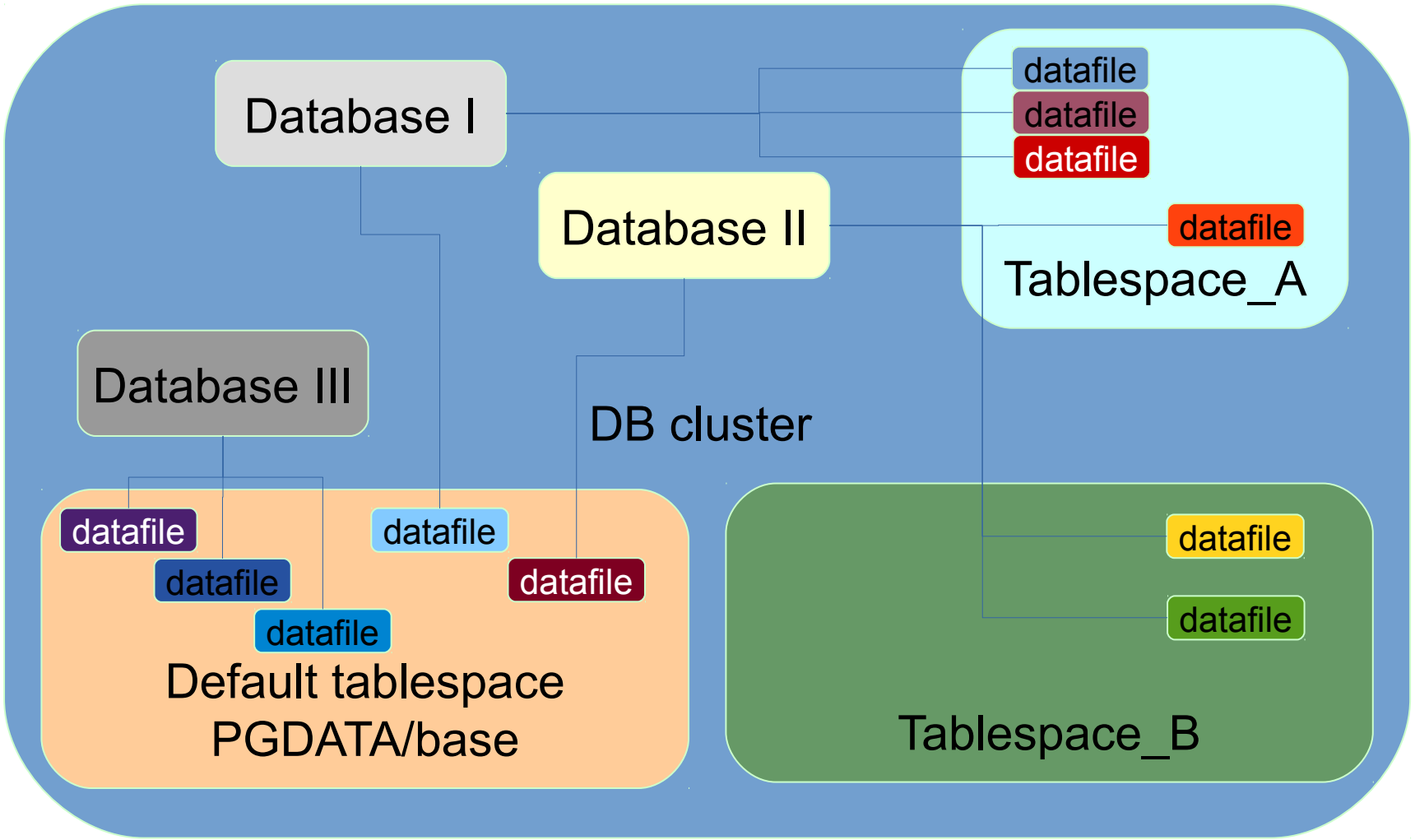
- PostgreSQL database structure (simplified)



Web: Postgres limits

Architecture – database structure

- PostgreSQL database structure (simplified)



Architecture - connections

PostgreSQL
The world's most advanced open source database.

Oracle

- Process named **listener** is responsible to handle new connections
  - listener.ora (network restrictions, TCP port)
  - sqlnet.ora (protocol configuration, kerberos...)
- Dedicated server processes per client
  - Usually used on U*NIX systems
- Multi-threaded server
  - Always used on Windows
  - Can be used on U*NIX systems as well

PostgreSQL

- Master process **postgres** (postmaster) listens for new connections
  - pg_hba.conf (user/database/network restrictions)
  - postgresql.conf (TCP port, kerberos, RDBMS cofiguration...)
- Dedicated server processes per client
  - Shared memory and semaphores are used for inter process synchronization
- Connection pooling by other products
  - PgBouncer
  - pgpool-II

Architecture notes

- Oracle
  - Decided that RDBMS is right and only place to manage database buffers
  - Promotes its ASM to have a direct control on file management (ASM is kind of LVM dedicated to Oracle)
- PostgreSQL
  - Relies on file system, OS file cache management
  - Do not re-implement features already implemented in OS, thus it use file system to store its data files (no RAW device support)

Security observations I

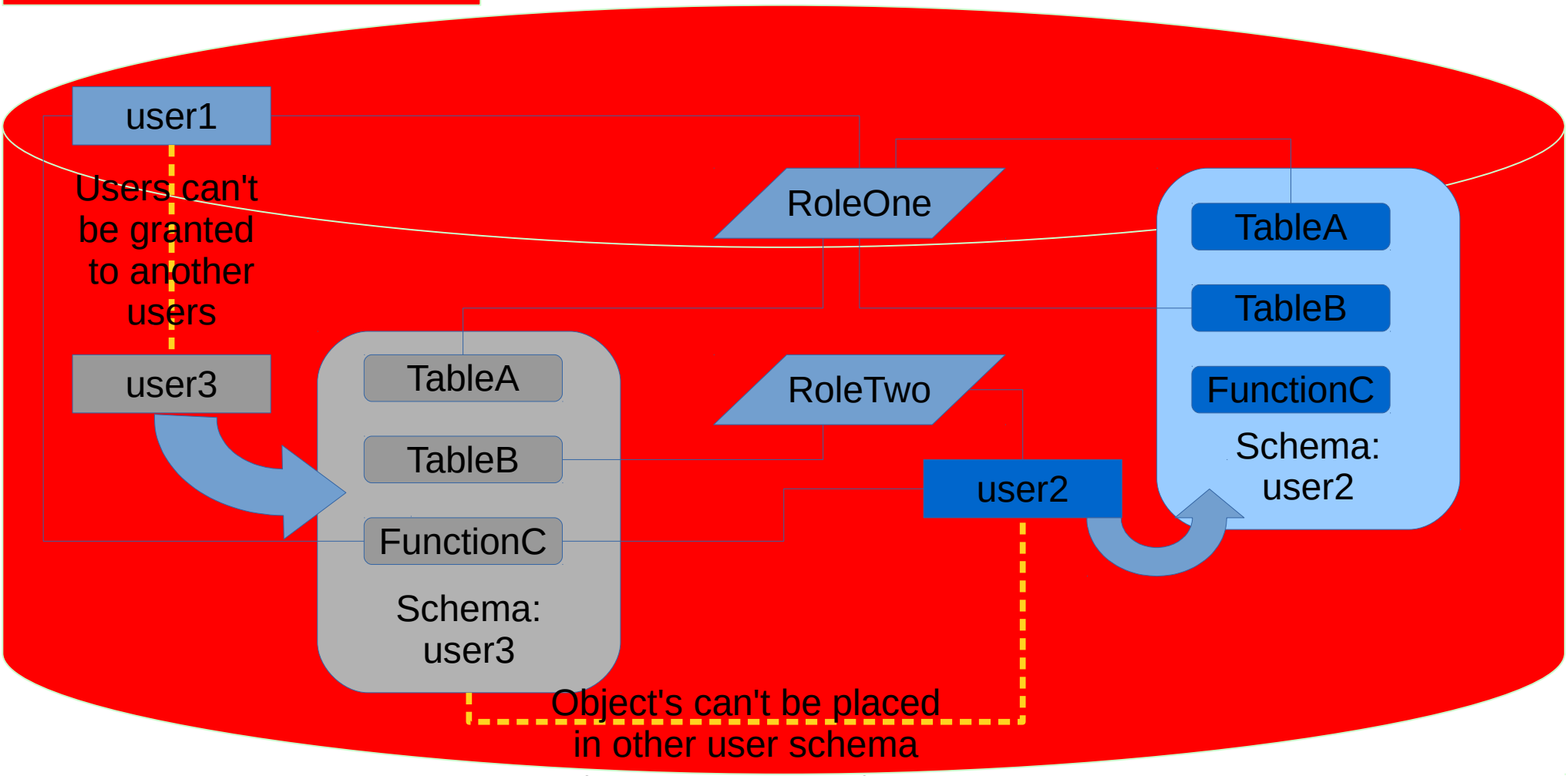
PostgreSQL
The world's most advanced open source database.

- Oracle has users and roles
  - Users and Roles are defined on DB level (not applies for PDB)
  - Users and Roles are different entities
- Postgres has **roles only**
  - Roles might be granted **`with login`** permission
- Oracle schema consist from a single user objects (schema = user)
  - Schema is not an object, so it can't be granted
- Postgres schema is a grantable (name-space) object
  - Ownership and usage on schema might be granted to roles
  - Objects owned by different roles (users) might reside within a single schema
  - Public schema might (and I think should) be dropped

Security observations I

PostgreSQL
The world's most advanced open source database.

SW installation:
/oracle/product/12.1.0.2/db_1
OS access control to files



Security observations II

PostgreSQL
The world's most advanced open source database.

- Oracle distinguish
  - System privileges (create table..., select any ...)
  - Object privileges (grant select on …)
- Postgres does not have such strong difference
  - Login permission is cluster wide kind of “system” privilege
  - Mostly all privileges are related to some objects including database object itself
    - Grant connect on database myDB
    - Grant usage on ...
    - Grant create on ...

Security observations III

PostgreSQL
The world's most advanced open source database.

- Oracle remote access control
  - IP address level: sqlnet.ora
    - tcp.validnode_checking = yes
    - tcp.invited_nodes = (hostname1, hostname2)
    - tcp.excluded_nodes = (192.168.10.3)
  - username password and create session is evaluated as next step
- Postgres
  - pg_hba.conf File
    - username/role_membership, database name, source IP address and authentification method is evaluated prior password validation
    - Password is evaluated as next step

Security observations IV

PostgreSQL
The world's most advanced open source database.

- Oracle procedures
  - **By default, all procedures are considered definer's rights**
  - Transactions can be controlled inside procedures
- Postgres functions
  - There are no procedures in Postgres
  - Function definition – **SECURITY INVOKER** indicates that the function is to be executed with the privileges of the user that calls it. That is the default
  - Transactions can't be started/commited/rolled-back within functions

Security features...

PostgreSQL
The world's most advanced open source database.

- ALL macro in grant commands
  - Expands to all at time of execution existing objects satisfying grant scope criteria
    - Grant execute on **ALL functions** in schema my_schema to …
- Alter default privileges
  - Does not affect existing objects, applied to newly created ones
  - Doc: ALTER DEFAULT PRIVILEGES

```
ALTER DEFAULT PRIVILEGES
    [ FOR { ROLE | USER } target_role
[, ...] ]
    [ IN SCHEMA schema_name [, ...] ]
    abbreviated_grant_or_revoke
```

Security fixes...

PostgreSQL
The world's most advanced open source database.

- Oracle
  - Bundle Patch Updates (BPUs) - a cumulative collection of fixes for a specific product or component
  - Patch Set Updates (PSUs) - a cumulative collection of high impact, low risk, and proven fixes for a specific product or component
  - Security Patch Updates - a cumulative collection of security bug fixes
  - Critical Patch Updates
  - Binaries are patches and SQL scripts to catalog are needed (catuppst.sql, catbundle.sql, datapatch.pl)
- Postgres
  - The PostgreSQL project aims to fully support a major release for five years.
  - PostgreSQL security updates are primarily made available as minor version upgrades.
    - Installation *usually* needs only stop the database server, install the updated binaries, and restart the server (check release notes before upgrade).

HA & DR

- OS clusterware (RHEL Pace Maker, PowerHA…)
  - Simply works if integrated storage replication is used
  - PCS is a bit tricky to configure well with PG built in replication, but possible
- Patroni HA - PostgreSQL HA with ZooKeeper, etcd or Consul
- There is nothing like Oracle RAC for PostgreSQL server implemented (and it seem there is no demand for that)
- Sharding (Postgres XL)
  - Oracle 12.2 seems to provide some support for sharding

Oracle replication

PostgreSQL
The world's most advanced open source database.

- Oracle DataGuard
  - Log shipping (log_archive_dest_n) by archiver
    - ARCHIVE_LAG_TARGET
  - Redo transmit by LGWR
    - ASYNC
    - SYNC
    - Delayed recovery - DELAY=minutes attribute of the LOG_ARCHIVE_DEST_n
  - Logical standby
  - Active Data Guard
  - Golden Gate

Postgres replication

PostgreSQL
The world's most advanced open source database.

- Postgres
  - Log-Shipping Standby Servers
    - archive_timeout
  - Streaming Replication
    - ASYNC (default)
    - SYNC - Synchronous Replication
    - Standby Server Settings recovery_min_apply_delay available from 9.4
  - Hot Standby (read only accessible standby)
  - Logical Standby
    - Slony, Bucardo
    - **pglogical** (faster than Slony – no triggers…) also nice for major release upgrades with minimum downtime – seems to be included in Postgres 10
  - BDR full multi-master capabilities, reusing much of the code from pglogical

Others…

PostgreSQL
The world's most advanced open source database.

- psql command line client
  - Comfortable interface, but be aware of default AUTOCOMMIT behavior
- --data-checksums initdb option
  - Page check-sums are calculated for all object in all databases in cluster
  - use pgbench to vefiry performance impact
  - Checksum is calculated on page read
  - Backup operate at file level, checksums are not calculated during backup
  - Available from 9.3 (09/2013) still not enabled by default :-(

Others… partitioning

PostgreSQL
The world's most advanced open source database.

- Postgres partitioning is implemented on top of inheritance feature
  - Declarative partitioning like in Oracle is not available
    - Improved version seems to be in Postgres 10 (no need to write triggers)
- Constraint on child tables
- Trigger on master table
  - Static IF… requires trigger compilation if new child partition tables are added
  - Trigger builds dynamic SQL – more overhead
- No global indexes on partitioned tables

Thanks for your attention

PostgreSQL
The world's most advanced open source database.

Q & A

/* takeaway slides continues */

Backup … and recovery

PostgreSQL
The world's most advanced open source database.

- Database [full or partial] dump
  - Oracle exp/imp, expdp/impdp
  - Postgres
    - pg_dump / pg_restore
      - “directory” format supports parallel dumps
    - pg_dumpall (use it for cluster globals only)
      - Load dump by call to psql
    - Thanks to MVCC, there is no “ORA-1555” risk during dump
      - For sure, the backup is consistent even if the database is used during the dump

Binary backups and recovery

PostgreSQL
The world's most advanced open source database.

- Offline! Works for Oracle, Postgres…
- Online Oracle database backups
  - Manual
    - Alter database (tablespace) begin backup, Copy corresponding datafiles, alter database (tablespace) end backup, store archived redologs needed for recovery
  - **Or use Oracle RMAN utility**
- Online Postgres cluster backup
  - Backup Control Functions
    - pg_start_backup(), pg_stop_backup(), same as above for Oracle [no TS level available]
  - pg_basebackup
    - Handle calls to backup control functions and might produce copy of postgres cluster or tar archive with the backup. Some features are available like tablespace mapping for convenient backup procedure/higher flexibility

Binary backups and recovery

PostgreSQL
The world's most advanced open source database.

- pgBarman
  - Some features similar to oracle RMAN
    - Recovery window / # of copies
    - Stores archived WALs together with Barman backups
    - Backup reports
    - Does not use “rman catalog”, backed up files with some barman metadata files are enough
    - Single backup might be archived to tape (tape integration is not part of pgBarman) – it disappears from backup reports, once retrieved from tape, pgBarman can use the backup again
    - RPO = 0 is possible with synchronous streaming replication (pg_receivexlog)
    - repmgr Integration might help to minimize RTO
- pgBackRest
  - More complicated configuration than Barman, incremental backups seems to be implemented slightly better
- pg_probackup promise page level incremental backups
  - I have not tested this yet...

Others… II

- Oracle dual table
  - `select function() from dual;`
  - SQL Loader, External tables
  - db links
- PostgreSQL
  - `select function();  select 5/8;`
  - `copy` command (client side, server side), file_fdw for CSV files, format compatible with COPY command required
  - Foreign Data Wrappers for many kinds of data sources, including Oracle database

Others… III

PostgreSQL
The world's most advanced open source database.

- Porting from Oracle PL/SQL
  - Oracle / Postgres – often similar, not always the same
    - ORA: trunc(date_variable, format)
    - PG: date_trunc('field', source)
- Pipelined functions are not implemented
- Group by can use column alias in postgreSQL

```
open2300db=> select date_trunc('hour', rec_datetime) as record_time,
   round(avg(temp_out), 2) as avg_temp,
   max(wind_speed_max) as max_wind_max
from open2300.weather where
   rec_datetime > now() - interval '3 hour'
group by record_time
order by record_time desc;
      record_time       | avg_temp | max_wind_max
------------------------+----------+--------------
 2016-02-15 22:00:00+01 |     3.04 |          2.4
 2016-02-15 21:00:00+01 |     3.23 |          2.5
 2016-02-15 20:00:00+01 |     3.66 |          2.3
 2016-02-15 19:00:00+01 |     4.11 |          2.6
```

Security observations V

PostgreSQL
The world's most advanced open source database.

- Oracle tablespace always belongs to a database
  - quotas might be used to limit tablespace usage by users
- Postgres tablespace is defined at cluster level
  - “create” on TS might be granted to a role
  - TS ownership to a role might be defined
  - There are no space usage quotas on tablespace, check FS free space
- Oracle database contains users defined inside DB, there is no database ownership concept
  - Grant scope is always within a database (PDB global users exception exists)
- Postgres database might be owned by a specific role
  - One role might have granted access on objects from multiple databases
  - Role attributes possible in scope of database – `alter role XXX set search_path = YYY,ZZZ in MY_DATABASE`

Security observations VI

PostgreSQL
The world's most advanced open source database.

- Oracle Advanced Security
  - Transparent Data Encryption
  - Kerberos (MS AD integration) is available without Advanced security as of 12.1 release, applies to older releases
  - Many other security features (VPD, RLS...)
- Postgres
  - SSO [Kerberos (MS AD integration)] is available
  - Row Security Policies are available with 9.5 release
  - TDE is not available (check Cybertec implementation)
  - Encryption is covered by separate module pgcrypto

Security observations VII

PostgreSQL
The world's most advanced open source database.

- Oracle [public] synonyms
  - Synonyms are used to reference another user (schema) objects
  - Might be defined as public – accessible to all users
- Postgres
  - `search_path` session environment variable is used to define scope of visible objects (schema search order), used similar to PATH in OS
    - Might be defined at cluster level
    - Users might have specified different search path values in particular databases

```
ALTER ROLE { role_specification | ALL } [ IN
DATABASE database_name ] SET configuration_parameter
{ TO | = } { value | DEFAULT }
```